

Center for  
Mathematical Economics  
Working Papers

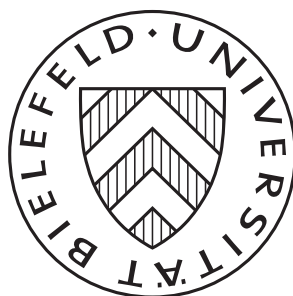
# 550

January 2016

## On Non-Cooperative Foundation and Implementation of the Nash Solution in Subgame Perfect Equilibrium via Rubinstein's Game

---

Papatya Duman and Walter Trockel



Center for Mathematical Economics (IMW)  
Bielefeld University  
Universitätsstraße 25  
D-33615 Bielefeld · Germany

e-mail: [imw@uni-bielefeld.de](mailto:imw@uni-bielefeld.de)  
<http://www.imw.uni-bielefeld.de/wp/>  
ISSN: 0931-6558

# On Non-Cooperative Foundation and Implementation of the Nash Solution in Subgame Perfect Equilibrium via Rubinstein's Game

Papatya Duman \*

Walter Trockel †

January 5, 2016

**JEL classification:** C7;C71;C72;C73

**Keywords:** Nash program; Non-cooperative foundation; Implementation; Nash solution; Rubinstein game; Subgame perfect equilibrium

## 1 Introduction

In the introduction of the their seminal article *The Nash Bargaining Solution in Economic Modelling*, Binmore et al. (1986) write:

*"This article studies the relation between the static axiomatic and the dynamic strategic approaches to bargaining. Its purpose is to clarify certain interpretive ambiguities in the axiomatic approach to provide a more solid grounding for applications of the Nash bargaining solution in economic modelling. The article may therefore be seen as a contribution to the "Nash program" as described by Binmore (1980)."*

In particular, it compares the Nash bargaining solution for cooperative two-person bargaining games introduced by Nash (1953) to the subgame perfect Nash equilibrium payoffs of the non-cooperative infinite horizon strategic two-person bargaining game with alternating offers due to Rubinstein (1982).

Binmore et al. (1986) comes out with an approximation of the Nash solution payoffs in two different cooperative bargaining games generated via different types of utility functions (time discounting versus risk averse von Neumann-Morgenstern) imposed on the basic dynamic model of Rubinstein (1982). There is not any known reasonable limit model when the discount factor or the probability of continuation of negotiations converges to 1. The infinite horizon model without discounting cannot be used to confirm the insights of the Rubinstein Theorem.

It is the aim of our present work to show that the support results for the Nash solution of Binmore (1980) or Binmore et al. (1986) which are only approximate can be modified in such a way as to establish an exact support. In a proposition and a theorem based on it, we shall provide such exact non-cooperative foundations for the Nash

---

\*Istanbul Bilgi University

†Murat Sertel Center of Advanced Economic Studies, Istanbul Bilgi University and Center of Mathematical Economics(IMW), Bielefeld University

bargaining solution, first via weakly subgame perfect, then via subgame perfect equilibrium. These results sharpen the "contributions to the Nash program" in Binmore (1980) and Binmore et al. (1986).

The Nash program that has its roots in some passages of Nash's work in the early 1950's is discussed in great details in Serrano (2004). Serrano writes in his introduction:

*"Similar to the microfoundations of the macroeconomics, which aim to bring closer the two branches of economic theory, the Nash program is an attempt to bridge the gap between the two counterparts of game theory (cooperative and non-cooperative). This is accomplished by investigating non-cooperative procedures that yield cooperative solutions as their equilibrium outcomes."*

He then quotes the following passage from Harsanyi (1974):

*"Nash (1953) has suggested that we can obtain a clear understanding of the alternative solution concepts proposed for cooperative games and can better identify and evaluate the assumptions to make about the players' bargaining behavior if we reconstruct them as equilibrium points in suitably defined bargaining games, treating the latter formally as non-cooperative games."*

As our modification of the Rubinstein game works for the diverse variants of the Rubinstein model with varying generality and complexity, we will work with a particularly simple and transparent special version that allows it to interpret the discount factor in both ways discussed in Binmore et al. (1986), namely as an indicator of either players' impatience or their risk aversion. Furthermore, we shall discuss the impact on implementability of the Nash solution of our exact and, in fact, also the approximate support results in Binmore (1980) and Binmore et al. (1986). This mechanism theoretical aspect is enormously relevant for applications of axiomatic bargaining solutions (for example, in wage bargaining) as discussed, for instance in Binmore et al. (1986) and in Gerber and Upmann (2006).

The relation between implementation theory and the Nash program has been extensively analyzed and discussed in Serrano (1997, 2004), Bergin and Duggan (1999), Trockel (2000, 2002a, 2002b, 2003).

In section 2, some basic notions of bargaining theory are introduced. Section 3 presents our version of the Rubinstein game. Section 4 with our Proposition on weakly subgame perfect support of the Nash solution is followed by the short section 5 on the concept of weakly subgame perfect equilibrium. In section 6, we establish in our Theorem a subgame perfect equilibrium support. Section 7 explains the impact of our two results on the problem of implementing the Nash solution in (weakly) subgame perfect equilibrium. In section 8, we collect concluding remarks.

## 2 Basic Concepts and Notation

We shall use the following two different types of games, namely two-person cooperative bargaining games and two-person non-cooperative games in extensive form, briefly extensive games. The definition of the latter ones is quite intricate though their illustrations via game trees are very intuitive. We shall use this notion as treated in Myerson (1991) (chapter 2) or in Mas-Colell et al. (1995)(chapter 7).

As to cooperative bargaining games, we shall use the following:

**Definition 1** *A two-person bargaining game is a pair  $(U, d)$  where  $d \in U \subset \mathbb{R}_+^2$  and  $U$  is non-empty, convex, compact and there exists an  $x \in U$  such that  $x \gg d$ . The set of two-person bargaining games is denoted by  $\mathbb{B}$ .*

**Definition 2** *A bargaining solution is a mapping*

$$\begin{aligned} L : \mathbb{B} &\longrightarrow \mathbb{R}^2 \\ (U, d) &\longmapsto L(U, d) \in U \end{aligned}$$

If we can associate any  $(U, d) \in \mathbb{B}$  with some extensive game  $G^{U, d}$  whose subgame perfect equilibrium payoff vectors coincide with  $L(U, d)$ , then the game  $G^{U, d}$  *supports* the solution  $L(U, d)$  of  $(U, d)$  by subgame perfect equilibrium. Such a support provides *an exact non-cooperative foundation* for the solution  $L$  in the sense of the Nash program (cf. Binmore et al. (1986), Serrano (2004)). Exact non-cooperative foundations for the Nash solution have been provided in van Damme (1986) [see also Naeve-Steinweg (1999) for a generalization], Howard (1992), Naeve (1999), Trockel (2000, 2002b).

In the present paper, we want to present an exact non-cooperative foundation for *the Nash Solution* based on the Rubinstein game.

The relevant notion of a subgame perfect Nash equilibrium due to Selten (1965) is defined as a Nash equilibrium on an extensive game which induces a Nash equilibrium on any subgame.

### 3 The Rubinstein Game

The Rubinstein infinite horizon strategic bargaining model with the two players' alternating offers is concerned with how to divide a unit of some perfectly divisible good with a resulting allocation for the two players. This game introduced by Rubinstein (1982) was meant to analyze "what 'will be' the agreed contract, assuming that both parties behave rationally". No link to axiomatic cooperative bargaining or even the Nash solution is indicated or, at least it appears so, intended. Discount factors  $\delta_1, \delta_2$  are assumed to be fix for both players. Possible consequences for the subgame perfect equilibrium regarding a relation to the Nash solution if  $\delta_1, \delta_2$  are close to 1 are not an issue. It was Binmore (1980) who related a dynamic version of Nash's simple demand game that he called "Modified Nash Demand Game II" to the (asymmetric) Nash solution, by approximating it by unique subgame perfect equilibrium payoff vectors of his strategic games where the discount factors  $\delta_1, \delta_2$  come close to 1. There are various versions of the original model of Rubinstein (1982) which has finite horizon predecessors in Ståhl (1972) and Krelle (1976). The most general version is used in Osborne and Rubinstein (1994) (chapter 7) where the set of feasible agreements  $X$  is a non-empty compact, connected set of some Euclidean space. In Binmore et al. (1986), in the introduction, the set  $X$  represents "physical outcomes" building the two players' "possible agreements". The formal model in their section 2 "strategic bargaining models" specifies this set as  $\bar{X} = \{x \in \mathbb{R}^2 | x_1 + x_2 \leq 1\}$ . In both cases, inefficient feasible agreements are principally possible. Rubinstein (1982) uses  $X = [0, 1]$  where the "split the pie" assumption excludes  $x_1 + x_2 < 1$  and implies efficiency of the agreements.

Binmore (1980) implicitly expresses the feasible set of alternatives as the set of payoff vectors in the utility image of some unspecified outcome set and explicitly assumes the set  $U$  of feasible payoff vectors to be a non-empty compact, convex set in  $\mathbb{R}^2$ . In this framework, the feasible proposal pairs  $(x_1, x_2)$  of the two players are payoff vectors and thus are directly comparable to the Nash solution point of  $(U, d)$ . Here  $d \in U$  is the disagreement payoff vector of the cooperative bargaining problem.

In the framework of the other approaches based on  $X$ , histories without any agreement at any time are mapped by the players' utility functions  $\pi_1$  and  $\pi_2$  onto such a  $d \in U$ , where  $U$  is  $\pi(X) = \{(\pi_1(x), \pi_2(x)) | x \in X\}$ .

As shown and in fact exploited in Binmore et al. (1986), the same underlying  $X$  may lead via different sets of players' utility functions to different sets  $U$  and accordingly different Nash solution points  $N(U, d)$ . Binmore et al. (1986) presents two detailed versions of strategic bargaining à la Rubinstein: one with time preferences and impatient players, the other one with exogenous risk of breakdown of negotiations with risk averse players who have von Neumann-Morgenstern utility functions. In both models, the subgame perfect equilibrium payoff vectors converge to the respective Nash solution points of the induced utility possibility set  $U$ , where  $\delta \in (0, 1)$  converges to 1.

We shall use a particularly simple and transparent version of the strategic bargaining that simultaneously allows both interpretations, namely impatience or risk aversion of players as represented by  $\delta$  as a discount factor as a probability of continuation of the negotiation process. This model is essentially that of Binmore (1980) and of Example 125.1 in Osborne and Rubinstein (1994). That will not affect the validity of our analysis for the more complex versions mentioned above. The interpretation of the discount factors we choose will have, however, a crucial impact on the application of our results to implementability in the mechanism theoretic sense. A reference for this model is also the collection of sections 6.7.1, 6.7.2 and 10.1.2 in Peters (2015).

Let  $X = [0, 1]$  represent the pie to be split among the two players. The payoff vector resulting from a division  $x = (x_1, x_2)$  with  $x_1 + x_2 = 1$  is determined by the utility functions  $u_i : X \rightarrow \mathbb{R}$  with  $u_i(x_i) = x_i$  for all  $i = 1, 2$ . Discounting utilities with  $\delta \in (0, 1)$  results in a payoff vector  $\delta^t x \in U$  for an agreement  $x \in X$  at time  $t \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ . We assume  $U = X$  and  $d = u(0, 0) = 0 \in U \subset \mathbb{R}^2$ .

For notational convenience, we shall identify players' proposals  $x_1, y_2 \in U$  with the payoff vectors  $(x_1, x_2) := (x_1, 1 - x_1)$  and  $(1 - y_2, y_2) \in \Delta := \{(z_1, z_2) | z_1 + z_2 = 1\}$ . This identification of  $\Delta$  with  $U$  via the two projections on the first and second coordinate, respectively, allows us to speak of proposals in  $U$  or in  $\Delta$  without creating confusion. We treat only the symmetric case with the discount factor  $\delta$  being the same for both players. The extension to the asymmetric case is possible like in the quoted literature and is straightforward.

In contrast to the finite horizon version of Ståhl (1972), in the Rubinstein game backward induction cannot be used for determining subgame perfect equilibria.

In our specific "split the pie" framework, there exists a "unique (not just essentially unique)" subgame perfect equilibrium (cf. Osborne and Rubinstein (1994)(p.125)). This unique subgame perfect Nash equilibrium  $\sigma_\delta^*$  is characterized as follows:

$$\sigma_{\delta,1}^* : \text{At } t \in 2\mathbb{N}_0, \text{ propose } x_\delta^* := (\frac{1}{1+\delta}, \frac{\delta}{1+\delta}); \text{ at } t \in 2\mathbb{N}_0 + 1, \text{ accept any proposal } z \in U \text{ of player 2} \\ \text{if and only if } z_1 \geq \delta \tilde{z}_{\delta,1}.$$

$$\sigma_{\delta,2}^* : \text{At } t \in 2\mathbb{N}_0 + 1, \text{ propose } y_\delta^* := (\frac{\delta}{1+\delta}, \frac{1}{1+\delta}); \text{ at } t \in 2\mathbb{N}_0, \text{ accept any proposal } z \in U \text{ of player 1} \\ \text{if and only if } z_2 \geq \delta y_{\delta,2}^*.$$

$x_\delta^*$  and  $y_\delta^*$  build the unique solution of the two equations  $x_2 = \delta y_2$ ,  $y_1 = \delta x_1$ , for  $x, y \in U$ .

The stationarity of these equilibrium strategies is a result rather than an assumption (cf. Osborne and Rubinstein (1994)(p.126)).

It can be easily verified that the Nash products  $x_{\delta,1}^* x_{\delta,2}^*$  and  $y_{\delta,1}^* y_{\delta,2}^*$  of  $x_\delta^*$  and  $y_\delta^*$  are the same. As both points are on the efficient boundary of  $U$ - this is also true in the more general case where the  $u_i$ 's are not identity functions; see for instance Figure 311.1 in Osborne and Rubinstein (1994)- with  $\delta$  converging to 1, both of  $x_\delta^*$  and  $y_\delta^*$  converge to  $z^* := N(U, 0)$ , the Nash solution point of  $(U, 0)$ .

The choice of  $X = [0, 1]$  like in Rubinstein (1982) used also in Example 120.1 of Osborne and Rubinstein (1994), is less natural than the  $\bar{X} := \{(x_1, x_2) \in \mathbb{R}_+^2 | x_1 + x_2 \leq 1\}$  chosen in Binmore et al. (1986), if one wants to compare the subgame perfect equilibrium payoffs with the Nash solution of cooperative games for classes considered usually in the literature. There  $U$  is generally a compact, convex (often strictly convex) set with  $d \in U$ . Our special case deals only with the efficient boundaries of such sets and  $d = (0, 0)$  does not satisfy  $d_1 + d_2 = 1$ . Making a proposal in our model corresponds to making a Pareto efficient proposal in the general case. In fact nothing relevant would change, if we replaced  $(U, 0) = (X, 0)$  by  $(\bar{X}, 0)$ .

## 4 An Exact Non-Cooperative Foundation

Denote the extensive form game of Rubinstein with discount factor  $\delta (= \delta_k) = \frac{k}{k+1}$  where  $k \in \mathbb{N}$ , described in the previous section, by  $G^k$  and its subgame perfect equilibrium payoff vector by  $\hat{z}^k$ ,  $k \in \mathbb{N}$ .

From the analysis in Binmore et al. (1986), we know that  $\lim_{k \rightarrow \infty} \hat{z}^k = z^*$ .

On this basis, we define now an extensive game  $G$  as follows:

At round 0, one of the two players of the bargaining game  $(U, d)$  is selected randomly with probability 1/2 to make a proposal  $z \in \Delta$ . After getting this offer, the other player can choose an element  $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ .

If she chooses 0, the proposal is accepted and the payoffs will be realized. If she chooses  $k \in \mathbb{N}$ , the proposal is rejected and the game  $G^k$  will be played, whose unique subgame perfect equilibrium payoff vector  $\hat{z}^k$  at  $t = 1$  results in the discounted payoff vector  $\delta_0 \hat{z}^k$  at  $t = 0$ . In order to simplify the notation, we assume, w.l.o.g.,  $\delta_0 := 1$ .

Like the Rubinstein games  $G^k$ ,  $k \in \mathbb{N}$ , the game  $G$  has an infinity of Nash equilibria, among them the one where every player always chooses proposal  $N(U, d)$  and always accepts this proposal and rejects any other one. Any other bargaining solution can be supported by Nash equilibria of  $G$  in an analogous way. This is essentially the situation we have in *Nash's simple demand game*.

Our final solution in the theorem will be based on subgame perfect equilibrium. But, as in contrast to the  $G^k$ ,  $k \in \mathbb{N}$ , the game  $G$  does not have any subgame perfect equilibrium, we shall first prove a proposition where we use the weaker concept of weak subgame perfectness due to Trockel (2011) that we are going to introduce next.

**Definition 3** *A Nash equilibrium of an extensive game is called **weakly subgame perfect** when it induces some Nash equilibrium in every subgame in which a Nash equilibrium exists.*

We shall very briefly discuss this concept in section 5. Here we will use it in order to state and prove our first non-cooperative support result.

**Proposition** For the bargaining game  $(U, d) = (U, 0)$ , the extensive game  $G(= G^{U,d})$  as defined above has an infinity of weakly subgame perfect equilibria with identical equilibrium path and equilibrium payoff vector  $z^* = N(U, d)$ .

**Proof** The proof is decomposed into several steps:

- One type of Nash equilibria is defined by the following rule for both players:

As the proposer choose  $z^*$ , as the follower accept exactly those proposals that are at least as good as  $z^*$ . In any Rubinstein subgame  $G^k$ , play according to the unique subgame perfect equilibrium.

It is obvious that no other proposal nor any other reaction to a proposal can constitute an advantageous unilateral deviation for any player. In these Nash equilibria,  $z^*$  is realized in the first round.

- There does not exist any subgame perfect equilibrium in this game.

This follows from the fact that  $G$  has subgames without Nash equilibria, namely those that directly follow any proposal  $z \in \Delta$  that offers to the other player a payoff smaller than her coordinate of  $z^*$ . Although this player should reject, there is no optimal  $k \in \mathbb{N}$  to do so.

- The Nash equilibria described in the first step are weakly subgame perfect.

The trivial subgame  $G$  has Nash equilibria according to the first step. Each of those induces on every Rubinstein subgame  $G^k$ ,  $k \in \mathbb{N}$ , a (subgame perfect) Nash equilibrium. All subgames starting at proposals that offer the other player a payoff at least as high as her coordinate of  $z^*$  have "acceptance" as the optimal, hence Nash equilibrium choice. Off the equilibrium path, these equilibria induce subgame perfect equilibria of Rubinstein games  $G^k$ ,  $k \in \mathbb{N}$ .

The only remaining subgames are those without equilibria described in the second step. So the Nash equilibria described in the first step are weakly subgame perfect.

- Any other Nash equilibrium fails to be weakly subgame perfect.

In order to establish this claim, consider a Nash equilibrium payoff vector  $\tilde{z} \neq z^*$ .

There are two possible ways how  $\tilde{z}$  may have been realized:

a) as an accepted proposal in the first round.

b) as the result of a subgame  $G^k$  that started right after a first proposal has been rejected.

W.l.o.g., let player 1 be the first proposer in the first round, and hence in every  $G^k$  if it is played after rejection of player 2.

Case (a): If player 1 proposes  $\tilde{z}$  with  $\tilde{z}_2 < z_2^*$  and player 2 accepts, this cannot possibly be a part of a Nash equilibrium, since player 2 could just reject with a sufficiently large  $k$  and ensure herself  $\hat{z}_2^k$  arbitrary close to  $N_2(U, d)$ , and  $\hat{z}_2^k > \tilde{z}_2$ .

If player 1 proposes  $\tilde{z}$  with  $\tilde{z}_2 > z_2^*$  and player 2 accepts, player 1 could have improved by proposing  $z^*$ -unless player 2 rejects  $z^*$ . So only if player 2's strategy contains rejection of  $z^*$ , then  $\tilde{z}$  could possibly be a Nash equilibrium payoff vector. But rejecting  $z^*$  is only possible via choosing some  $G^k$ ,  $k \in \mathbb{N}$ . As in each  $G^k$ ,

the unique subgame perfect equilibrium payoff for player 2 would be  $\hat{z}_2^k < z_2^* < \tilde{z}_2$ , none of them would justify rejection of  $z^*$ . Therefore, the payoff vector  $\tilde{z}$  can possibly result only from a Nash equilibrium that is not weakly subgame perfect and satisfies  $\delta_0 \tilde{z}_2 = \tilde{z}_2 > \hat{z}_2^k = \delta_0 \hat{z}_2^k$ .

Case (b): Suppose  $\tilde{z}$  is a weakly subgame perfect equilibrium payoff vector of  $G$ . Then  $\tilde{z} = \delta_0 \hat{z}^k = \hat{z}^k$  for some  $k \in \mathbb{N}$ . But then player 2 could improve by choosing  $k + 1$  with discounted payoff vector  $\delta_0 \hat{z}^{k+1} = \hat{z}^{k+1}$ . This implies that  $\tilde{z}$  is not a Nash equilibrium payoff vector of  $G$ , a contradiction.  $\square$

## 5 Remarks on Weakly Subgame Perfect Equilibria

In contrast to the games  $G^k$ ,  $k \in \mathbb{N}$ , the game  $G$  has an infinity of weakly subgame perfect equilibria. How bad is this? There is no coordination problem involved as long as both players stay on the equilibrium path and the equilibrium payoff is uniquely determined. So the multiplicity of those equilibria appears to be harmless, in particular as subgame perfect equilibria also may have multiple ways of behavior off the equilibrium path.

So the criticism could only be based on the lack of credible threats to reject, because there is no optimal way of rejecting! But from a decision theoretical point of view, this criticism is dubious. If there is a choice between money amounts  $\{-50, 1, 2, \dots, 10\}$ , we take it for granted that -50 is rejected (via accepting 10). If the choice is among  $\{-50\} \cup \mathbb{N}$ , do we think that -50 will be accepted just because there is no best alternative? In real life, we avoid very bad or worst cases even if we are unable to do that in an optimal way.

But as this is a controversial point, we will provide a modified non-cooperative support result via subgame perfect equilibria in the next section.

## 6 Subgame Perfect Exact Support

When attempting to base an exact subgame perfect equilibrium foundation for the Nash solution based on Rubinstein's games  $G^k$ ,  $k \in \mathbb{N}$ , the dilemma is the appearance of those subgames starting right after an initial proposal that do not have any Nash equilibrium. There are two potential ways out, one may consider:

1. Add a best alternative to the set  $\mathbb{N}_0$ . We did not see any natural way to do so. We might end up with no or a multiplicity of subgame perfect equilibria. Anyway, we did not follow this approach.
2. Stop those subgames starting right after the first proposals from being subgames. In order to do so, we modify our original game  $G$  in the following way: in the beginning, the proposer is chosen randomly with probability  $1/2$ , but both players do not observe that random choice. So each player has probability  $1/2$  that she is the chosen proposer and  $1/2$  that she has to react to her opponent's proposal. Accordingly, both players' strategies have to contain full descriptions of what they would propose and how they would react to any possible proposal  $z \in \Delta$ .

We follow that approach!



In the beginning, the referee throws a fair coin in order to decide who of the two players starts with a proposal  $z \in \Delta$ . But the result of this random choice is not observed by the players. Then, not knowing whether they will start with a proposal or a rejection to the proposal, but knowing that any Rubinstein game played after a rejection starts with another proposal of the proposer, both players simultaneously and independently choose a pair consisting of a proposal and a reaction function on the set of possible proposals.

On this basis, we define now an extensive game  $\tilde{G}$  as follows:

At stage 0, the following things happen:

1. The referee throws a fair coin in order to randomly but privately determine which player will act as the proposer.
2. Both players, not knowing which of them will act as the proposer, simultaneously submit pairs  $(x_1, f_1), (x_2, f_2) \in U \times \mathbb{N}_0^U$  to the referee, where  $x_i$  are their proposals,  $f_i$  are reaction function to their opponents proposal, for  $i = 1, 2$ .
3. The referee informs both players on their randomly determined roles. The proposer whose role is now common knowledge is w.l.o.g. called player 1.

At stage 1, the game either ends if  $f_2(x_1) = 0$  or continues to stage 2 if  $f_2(x_1) = k \in N$ . In this case, player 1 starts at stage 2 with a proposal  $x_1 \in U$  in the Rubinstein game  $G^k$ . The rest of the game is just playing this Rubinstein game  $G^k$ .

The specific structure of  $\tilde{G}$  at stage 0 has two consequences that are crucial for our Theorem:

First, it guarantees that the players' choices of proposals together with their reaction functions build actions at  $t = 0$  in non-singleton information sets. Therefore, no subgame starts with such actions. Secondly, the full information of both players about their respective roles as proposer and reactor before the start of actions at stage 2 prevents the annoying situation that all  $G^k, k \in N$  would start only at non-singleton information sets. In that case, the only subgame of  $\tilde{G}$  would be  $\tilde{G}$  itself.

Clearly,  $N(U, d)$  would still be a subgame perfect equilibrium payoff vector. But the whole plethora of Nash equilibria would become subgame perfect ones, too!

Under the aspect of trying to support the Nash solution, we would essentially be back at Nash's simple demand game. Notice that in  $\tilde{G}$ , the Rubinstein subgames  $G^k$  of  $G, k \in N$  will reappear twice : once via a rejection of player 2 and once as a rejection of player 1 in that part of  $\tilde{G}$  that follows the non-realized choice of player 2 as the proposer. Only those  $G^k, k \in N$  following a rejection by player 2 are potentially effective for the outcome of the game. But also in the other (now irrelevant) Rubinstein games, following rejections by player 1 of proposals by player 2, subgame perfectness of  $\tilde{G}$  requires the players to play subgame perfect equilibria.

It is impossible to just cancel that at stage 1 irrelevant part of the game tree as it is relevant for the players' choices at stage 0 before they are informed about their respective later roles.

This modification of our original game  $G$  is illustrated in the following two figures:

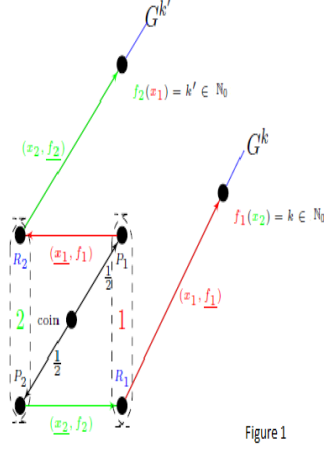


Figure 1

Figure 1 and 2 are equivalent stylized illustrations of stages 0 and 1 of  $\tilde{G}$ .

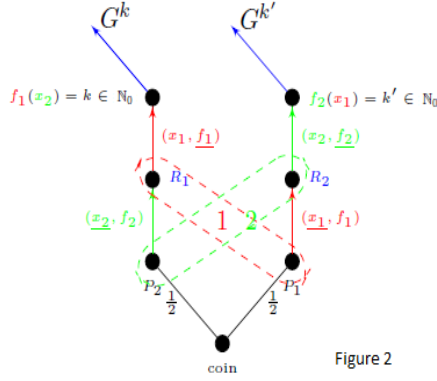


Figure 2

Notice that these figures are only schematic illustrations of the game  $\tilde{G}$  in its first stages  $t = 0, 1$  rather than complete game trees. The infinite action sets for both players are represented in these figures only by one typical action for each player, namely  $(x_1, f_1)$ ,  $(x_2, f_2)$ . In the Figures  $G^0$  is the degenerate game consisting of the singleton set  $\{0\}$  representing acceptance of a proposal.

The construction in Figure 2 is similar to the way in which Sudhölter et al. (2000) define the canonical extensive form for the Battle of Sexes game. It has precisely the intended effect in our present context. The one-player subgames without optimal actions in the reduced game have vanished now.

The only remaining subgames of the modified game  $\tilde{G}$  are  $\tilde{G}$  itself and the Rubinstein games  $G^k$ ,  $k \in \mathbb{N}$ .

In  $\tilde{G}$ , any subgame perfect equilibrium is a pair  $(x_1, f_{1,k^1}; x_2, f_{2,k^2})$  with  $(x_1, x_2) = z^*$  and  $f_{i,k^i} : U \rightarrow \mathbb{N}_0$  such that

$$f_{i,k^i}(x) = \begin{cases} 0 & : x_i \geq z_i^* \\ k^i & : x_i < z_i^* \end{cases}$$

and  $k^i \in \{k \in \mathbb{N} | \hat{z}_i^k > x_i\}$ ,  $i = 1, 2$ .

After the first moves of both players, the game ends either in its equilibrium with payoff vector  $z^*$  or, else, in some game  $G^k$ ,  $k \in \mathbb{N}$ , where the unique subgame perfect equilibrium is induced. The multiplicity of subgame perfect equilibria arises from the various  $k^i$  that may be chosen for  $f_{i,k^i}$ ,  $i = 1, 2$ . But the equilibrium path is unique.

We can formulate the modified version of our support result now as follows:

**Theorem** *For the bargaining game  $(U, d) = (U, 0)$ , the extensive game  $\tilde{G}(= \tilde{G}^U)$  as defined above has an infinity of subgame perfect equilibria with identical equilibrium path and equilibrium payoff vector  $z^*$ .*

**Proof** *This theorem is in fact a corollary to the Proposition.*

*All Rubinstein games  $G_k$ ,  $k \in \mathbb{N}$  occur as subgames in  $\tilde{G}$ . There is no subgame anymore however, that starts after a proposal with actions  $k \in \mathbb{N}_0$ . The  $k$  chosen after a proposal  $x_1$  of player 1 is now determined by the simultaneous choice  $(x_2, f_2)$  of player 2 at both points of her information set at  $t = 0$  via  $k := f_2(x_1)$ .  $k' := f_1(x_2)$  has been eliminated from further consideration by the referee's random choice of player 1 as the proposer. Accordingly, a lack of an optimal way of rejecting a proposal cannot destroy the subgame perfectness of a weakly subgame perfect Nash equilibrium. As at  $t = 0$ , there do not exist singleton information sets for the players, the games  $\tilde{G}$  and  $G_k$ ,  $k \in \mathbb{N}$  build all subgames of  $\tilde{G}$ .*

*Now, we define a Nash equilibrium of  $\tilde{G}$  literally as in the first step of the proof of our Proposition. Any weakly subgame perfect equilibrium of  $G$  "becomes" (corresponds to) subgame perfect equilibria of  $\tilde{G}$ . Any other Nash equilibrium payoff vector could in  $G$  only result from non weakly subgame perfect behavior in some  $G^k$ , thus in  $\tilde{G}$  only from violating subgame perfectness in that  $G^k$ . This proves our Theorem.  $\square$*

**Remark** *The subgame perfect Nash equilibria of  $\tilde{G}$  are in fact even sequential Nash equilibria! The beliefs of both players can be expressed by probabilities on their non-singleton information sets. Whatever probabilities there may be, however, they do not have any influence on their decisions at these information sets, as every choice  $(x_i, f_i)$ ,  $i = 1, 2$  is intended to be optimal at each point of the information set, respectively. Only the relevant parts of those pairs  $(x_i, f_i)$ ,  $i = 1, 2$  are different at the different points of the information sets.*

## 7 From Support to Implementation

The step from a non-cooperative support of a cooperative solution to a mechanism theoretic implementation is not trivial and requires some care and in fact certain assumptions.

In Serrano (1997), we find the following passage:

*"The Nash Program and the abstract theory of implementation are often regarded as unrelated research agendas. Indeed, their goals are quite different: while the former attempts to gain additional support for cooperative solutions based on the specification of certain non-cooperative games, the latter tries to help an incompletely informed designer implement certain desirable outcomes. However, it is misleading to think that their methodologies cannot be reconciled. A common criticism that is raised against the mechanisms in the Nash program is that they are not performing real "implementations" since their rules depend on the data of the underlying problem (say the characteristic function) that the designer is not supposed to know."*

Bergin and Duggan (1999) also emphasizes the importance of independence of the game rules expressed by a mechanism of the players' preference profiles. Serrano (2004) (p.1) writes:

*"It is true that many contributions in the Nash program write down a mechanism whose rules depend on the characteristic function, but those mechanisms become more useful if their rules can be adapted in terms of outcomes, thereby making them independent of preferences. In particular, this means that if the characteristic function is known to the designer and the only goal of the Nash program were seen as suggesting a non-cooperative game whose set of equilibrium payoffs is  $F(N, V)$ , one can achieve this by the trivial mechanism: if all players agree on a given payoff profile in the solution, implement it; otherwise implement a bad outcome. This point, which was in the folklore of the practitioners of the Nash program for a while, was fleshed out by Bergin and Duggan (1999)(Proposition 1). However this mechanism is not useful and does not shed any light on the meaning of the specific solution (as such, the mechanism "works" for any solution)."*

The rules for implementation described here make the game a *nonsensible* one in the sense of Nash. They do not provide any insight about specific properties of the specific cooperative solution under consideration.

The very fact that the presence of an outcome space is an additional ingredient in mechanism theory as compared to the Nash program indicates that these two can hardly be considered "equivalent", as claimed, for instance in Dagan and Serrano (1998)(abstract). Detailed treatments of the relation between the Nash program and implementation theory can be found in Bergin and Duggan (1999), Trockel (2000, 2002a) and Serrano (2004).

The conditions that are necessary in order to have a non-cooperative support that automatically provides an implementation in some equilibrium concepts are often satisfied in models in the literature (cf. Moulin(1984), Howard (1992)). However, strictly speaking, there *solution based social choice rules* are implemented.(cf. Trockel (2003)).

This holds also true in principle for the model of Rubinstein (1982), as used in Binmore et al. (1986) and Osborne and Rubinstein (1994). However, the situation there is different. In a strict sense, these results are not to provide a non-cooperative implementation for the Nash solution on a given prespecified class of two person cooperative bargaining games. They rather just define such an implementation for the classes of those bargaining games generated by their game forms together with their different types of utility functions. From a puristic point of view, there is a missing axiomatization of the Nash solution on those classes. Clearly, this solution is still well defined as the maximizer of the Nash product. In fact, the section 3 in Binmore et al. (1986) has the heading *Nash solution as an approximation to the equilibria*. This terminology differs from the one prevailing in the literature on non-cooperative foundation where one thinks of supporting the Nash solution by equilibria of non-cooperative games, exactly or by approximation. In fact, the equilibria do approximate the Nash solution.

The Nash program is based on two passages in Nash (1951) and Nash (1953). While the first one may be interpreted as giving a higher priority to the strategic than to the axiomatic approach to bargaining, the second one from the introduction in Nash (1953) emphasizes the equal importance of both approaches:

*"We give two independent derivations of our solution of the two-person cooperative game. In the first, the cooperative game is reduced to a non-cooperative game. To do this, one makes the players' steps of negotiation in the cooperative game become moves in the non-cooperative model. Of course, one cannot represent all possible bargaining devices as moves in the non-cooperative game. The negotiation process must be formalized and restricted, but in such a way*

that each participant is still able to utilize all the essential strengths of his position.

The second approach is by the axiomatic method. One states as axioms several properties that it would seem natural for the solution to have and then one discovers that the axioms actually determine the solution uniquely. The two approaches to the problem, via the negotiation model or via the axioms, are complementary; each helps to justify and clarify the other.”

A *justification* in Nash’s sense of the Nash bargaining solution concept on a specified whole class of bargaining games via the non-cooperative approach as opposed to just as the Nash solution point of one specific bargaining game requires intuitively that each game in that class can be generated or represented by a game in a family of “similar” strategic games. Such kind of *uniform* support for a whole class leads naturally to a common game form underlying that family of strategic games. That is an important step towards implementation.

Binmore et al. (1986) by using the terms *time-preference Nash solution* and *von Neumann-Morgenstern Nash solution* stress the fact that their two models provide approximate supports for the Nash solution on different sets of bargaining games.

The other direction of the Nash program, namely *justification* and *clarification* of the Rubinstein approach, is not so obvious. This model is defined with discount factors as important ingredients, technically and conceptually. But as soon as the discounting is taken serious rather than almost neglected the Nash solution of the induced cooperative bargaining games may differ significantly from the subgame perfect equilibrium payoff vectors of the strategic games. As far as *implementation* is concerned, this fact is not disturbing, however. What in implementation theory has to be implemented is a social choice rule. And the implementation is not conceived as a *justification* or *clarification* like in the Nash program. Rather the social choice rule is the inherently justified solution method for a social choice problem with a specified outcome set. Its implementation in a strategic equilibrium concept represents the idea of realizing something already accepted as socially desirable via strategic interaction in the society according to certain rules, namely a *mechanism* or *game form*.

In situations where solutions of certain games can be identified with social choice rules, or as Hurwicz (1994) termed them, *desirability correspondences*, non-cooperative foundation may extend to implementation.

In our context, the Nash solution has to play the role of the social choice rule. One can easily see that the search for a natural and adequate outcome set needed for implementation leads to different results for the various versions of sequential bargaining in Binmore (1980), Binmore et al. (1986) or Osborne and Rubinstein (1994). Consequently, the generated utility spaces or cooperative bargaining games may be quite different.

In this context, it is important to notice that our use of the payoff set  $U$  (identified with the underlying  $X = [0, 1]$ ) in this paper on which the negotiation is modeled rather than on the set  $\bar{X} = \{x \in \mathbb{R}_+^2 \mid x_1 + x_2 \leq 1\}$  is just for convenience and simplicity of presentation, like the treatment in Binmore (1980). It could as well have been formulated via the models in Binmore et al. (1986) like for instance in Osborne and Rubinstein (1994)(Proposition 310.3).

From the conceptual point of view, there is a fundamental difference between the two models in Binmore et al. (1986) and therefore also in the two interpretations of the factor  $\delta$  in the results of our present paper as far as the implementability problem is concerned.

If the Nash solution should be implemented on a class of bargaining games resulting from players who are really impatient then the time preference Nash solution cannot even approximately be implemented because the discounting factor cannot be made close to 1 by the designer. So the time preference Rubinstein approach could lead to implementation of the Nash solution only on a class of bargaining games generated by the strategic games with players whose idiosyncratic discount factors would have to be close enough to 1. Moreover, because of symmetry of the Nash solution, these discount factors would have to be the same for any pair of players negotiating. That means that practically we can forget about approximate implementation in that context.

A similar argumentation holds for those bargaining games in the second model of Binmore et al. (1986) that are generated when players are strongly risk averse. Only uniform weak enough risk aversion would allow implementation of the Nash solution for that induced class of bargaining games.

A third possibility is to think of a pool of risk neutral perfectly patient players. In that case, the second model of Binmore et al. (1986) allows to give the breakdown probabilities as instruments into the hands of the designer. And only then we would get arbitrarily close to implementation of the Nash solution on the induced class of cooperative bargaining games. Only in that case our two present results could be transformed into conceptually meaningful implementation results. Yet, this domain of the Nash social rule would be very specific and small.

## 8 Concluding Remarks

Our exact non-cooperative foundation results for the Nash bargaining solution that we have presented in this article are, as we mentioned earlier, not the only ones in the literature. Earlier such results in various different ways may be found in van Damme (1986), Howard (1992), Naeve-Steinweg (1999), Naeve (1999) and Trockel (2000, 2002b). For an approximate support via a modification of Rubinstein's game that, differently from Rubinstein, works also for  $n > 2$ , (see Moulin (1984)).

As to the tasks of *clarification* and *justification*, it is interesting to point out to the similarity of insights into the meaning of the Nash solution provided by the Rubinstein alternating offer approach and by the non-cooperative strategic unique Nash equilibrium support and implementation in Trockel (2000) based on Walrasian payoff functions. In the latter one, this Walrasian property of the Nash solution (cf. Trockel (1996)) represents perfect competition that simulates an abundance of *outside options* for both players in a game protecting them from exploitation by their respective opponents. In the Rubinstein game with almost negligible discounting and in our modified games, it is the *infinity of future options*, (almost) equally valuable, that creates the same effect. That suggests the interpretation of implementing the Nash solution as a sort of surrogate for sufficient competitive pressure.

If the implementation is not an issue but only a non-cooperative foundation that helps to *justify* the Nash bargaining solution and to *clarify* its meaning then our results will work equally well (but as we think not better) as the approximate results based on the Rubinstein game. The advantage of our modified Rubinstein game forms lies in the fact that they somehow make the subgame perfect equilibrium payoff function continuous at  $\delta = 1$ .

While the Rubinstein games with  $\delta$  converging to 1 induce a limit for the associated sequence of subgame perfect equilibrium payoffs, there is no associate *limit model* whose subgame perfect payoff equilibrium vectors would confirm

that result. This can be changed by modifying the Rubinstein game without changing the unique subgame perfect equilibrium payoff vector in the following manner: in the case of discount factor  $\delta$ , we take almost the same game tree as in our game  $G$  and change the payoffs of the game as follows:

After the first proposal of ,say, player 1, player 2 has the choice either to choose a number from the set  $\{1, 2, \dots, K\}$  where  $K$  is the largest  $k$  for which  $\frac{k}{k+1}$  is smaller than  $\delta$ , in which case the payoff vector is that of the subgame perfect equilibria of the Rubinstein game  $G^k$  or to accept via  $k = 0$ . The unique subgame perfect equilibrium payoff vector of this game is that of the Rubinstein game  $G^K$ . In any subgame perfect equilibrium, the proposal is that equilibrium proposal and exactly those proposals are accepted that are at least as good (for the accepting player) as that equilibrium proposal. In all other situations, the choices are determined by the subgame equilibria of the subgames,  $G^k$ ,  $k = 1, \dots, K$ .

The sequence of games  $\widehat{G}^K$ ,  $K \in \mathbb{N}$  defined in this way converges to our game  $G$  for  $K$  growing to infinity. Each of the games in the sequence has the same subgame perfect equilibrium payoffs as the corresponding Rubinstein games  $G^K$ . Their limit is the weakly subgame perfect equilibrium payoff vector of  $G$ . As for the finite games in our sequence, weakly subgame perfect equilibria and subgame perfect equilibria coincide, we get continuity at infinity of the weakly subgame perfect equilibrium payoff vector function.

**Remark** *Totally analogous results to our Proposition and our Theorem can be proved via using Ståhl's rather than Rubinstein's model. In the games  $G^k$ ,  $k$  would have been to be replaced by a double index  $(k, l)$ , where  $k$  represents the discount factor  $\delta$ , and  $l$  represents the number of stages of a (finite horizon!) Ståhl game.*

## References

- Bergin, James, and Duggan, John. *An implementation-theoretic approach to non-cooperative foundations*. 1999: Journal of Economic Theory 86.1: 50-76.
- Dagan, Nir, and Serrano, Roberto. *Invariance and randomness in the Nash program for coalitional games*. 1998: Economics Letters 58.1: 43-49.
- van Damme, Eric. *The Nash bargaining solution is optimal*. 1986: Journal of Economic Theory 38.1: 78-100.
- Binmore, Ken. *The Nash bargaining theory II*. 1987: Chapte 4 in Binmore and Dasgupta (eds): The Economics of Bargaining, Oxford, Basil Blackwell : 61-76.
- Binmore, Ken, Rubinstein, Ariel, and Wolinsky, Asher. *The Nash bargaining solution in economic modelling*. 1986: The RAND Journal of Economics: 176-188.
- Gerber, Anke, and Upmann, Thorsten. *Bargaining solutions at work: Qualitative differences in policy implications*. 2006: Mathematical Social Sciences 52.2: 162-175.
- Harsanyi, John C. *An equilibrium-point interpretation of stable sets and a proposed alternative definition*. 1974: Management Science 20.11 : 1472-1495.

- Howard, John V. *A social choice rule and its implementation in perfect equilibrium*. 1992: Journal of Economic Theory 56.1: 142-159.
- Hurwicz, Leonid. *Economic design, adjustment processes, mechanisms, and institutions*. 1994: Review of Economic Design 1.1: 1-14.
- Krelle, Wilhelm. *Preistheorie (Part 2)*. 1976: Tübingen: JCB Mohr (Paul Siebeck).
- Mas-Colell, Andreu, Whinston, Michael Dennis and Green, Jerry R. *Microeconomic Theory. Vol. 1*. 1995: New York: Oxford University press.
- Moulin, Herve. *Implementing the Kalai-Smorodinsky bargaining solution*. 1984: Journal of Economic Theory 33.1: 32-45.
- Myerson, Roger B. *Game theory: Analysis of Conflict*. 1991: Harvard University.
- Nash, John. *Non-cooperative games*. 1951: Annals of Mathematics: 286-295.
- Nash, John. *Two-person cooperative games*. 1953: Econometrica: 128-140.
- Naeve, Jörg. *Nash implementation of the Nash bargaining solution using intuitive message spaces*. 1999: Economics Letters 62.1: 23-28.
- Naeve-Steinweg, Elisabeth. *A note on van Damme's mechanism*. 1999: Review of Economic Design 4.2: 179-187.
- Osborne, Martin J., and Rubinstein, Ariel. *A Course in Game Theory*. MIT press.
- Peters, Hans. *Game Theory. A Multi-Leveled Approach, Second Edition* 2015: Springer-Verlag, Berlin Heidelberg.
- Rubinstein, Ariel. *Perfect equilibrium in a bargaining model*. 1982: Econometrica: 97-109.
- Selten, Reinhard. *Pieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit: Teil 1: Bestimmung des dynamischen Preisgleichgewichts*. 1965: Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics: 301-324.
- Serrano, Roberto. *A comment on the Nash program and the theory of implementation*. 1997: Economics Letters 55.2: 203-208.
- Serrano, Roberto. *Fifty years of the Nash program, 1953-2003*. 2004: Investigaciones Economicas 29: 219-258.
- Ståhl, Ingolf. *Bargaining theory*. 1972: Economics Research Institute at the Stockholm School of Economics.
- Sudhölter, Peter., Peleg, Bezalel. and Rosenmüller, Joachim. *The canonical extensive form of a game form: Part II-Representation*. 2000: Journal of Mathematical Economics 33: 299-338.
- Trockel, Walter. *A Walrasian approach to bargaining games*. 1996: Economics Letters 51.3: 295-301.



- Trockel, Walter. *Implementations of the Nash solution based on its Walrasian characterization*. 2000: Economic Theory 16.2: 277-294.
- Trockel, Walter. *Integrating the Nash program into mechanism theory*. 2002a: Review of Economic Design 7.1: 27-43.
- Trockel, Walter. *A universal meta bargaining implementation of the Nash solution*. 2002b: Social Choice and Welfare 19.3: 581-586.
- Trockel, Walter. *Can and should the Nash program be looked at as a part of mechanism theory?*. 2003: Murat r. Sertel and Semih Koray (eds): Advances in Economic Design, Springer-Verlag, Berlin Heidelberg: 153-174.
- Trockel, Walter. *An exact non-cooperative support for the sequential Raiffa solution*. 2011: Journal of Mathematical Economics 47(1): 77-83.